

Accession For

NTIS CSASI
DTIC TAS S
Unamerated S
Juntified than 2



1.

Final Scientific Report

Grant Number AFOSR-76-2927

t Number AF



(A) Research Completed

- 1. Augmenting Shapiro-Wilk test for normality. The expected of the i-th order statistic is considered to be $\gamma_0+\gamma_1c_i+\gamma_2(c_i^2-\lambda)+\gamma_3(c_i^3-\mu c_i)+\ldots$, where c_i is the expected value of the i=th order statistic from the N(0,1) distribution. If the underlying distribution is normal, then $\gamma_1=\sigma$, $\gamma_2=0$, $\gamma_3=0$, Shapiro-Wilk test is designed to test the hypothesis $\gamma_1=\sigma$. The possibility of testing the multitiple hypothesis $\gamma_1=\sigma$, $\gamma_2=0$, $\gamma_3=0$, ..., is considered in the present paper and improvements in efficiency over the Shapiro-Wilk test are demonstrated.
- 2. On linear combination of order statistics. Two short and elementary proofs for the necessary and sufficient conditions for the asymptotic normality of linear combinations of order statistics (U_i) from the rectangular distribution U(0,1) over [0,1] are given. The first one based on the corresponding simple result for random samples from the exponential distribution, uses the probability integral transformation and linearizes it by the Taylor formula applied at a random expansion point. The latter provision permits small linear terms which in turn cause the sums of the remainder terms to tend stochastically to zero. The second proof is based on the representation of U_in as a ratio of partial sums of i.i.d. exponential r.v.'s and then again refers to the above cited corresponding result for exponential order statistics.
- 3. Problems of association for bivariate circular data and new test of independence. The question of association for bivariate circular data is introduced by giving a thorough summary of previous work in the area. This includes not only attempts at defining notions of circular correlation but also circular regression and testing dependence, from bivariate random sample on the torus. Also new test of independence which is designed specifically for axially symmetric type data is proposed. The proposed statistic is motivated as a measure of discrepancy between what is observed and what is "expected" under independence.

Appropriation unlimited.

THE FILE COP

Its computational form is given in terms of circular gaps, and its distribution under the null hypothesis of independence is derived.

- 4. Asymptotically distribution-free aligned rank order tests for composite hypotheses for general linear models. For general multivariate linear models, a composite hypothesis does not usually induce invariance of the joint distribution under appropriate groups of transformations, so that genuinely distribution-free tests do not usually exist. For this purpose, some aligned rank order statistics are incorporated in the proposal and study of a class of asymptotically distribution-free tests. Tests for the parallelism of several multiple regression surfaces are also considered. Finally, the optimal properties of these tests are also discussed.
- 5. Convergence and remainder terms in linear rank statistics. In this paper, a new approach to the asymptotic normality of linear rank statistics for the regression case is provided under general alternatives. The problem of the rate of convergence of the statistics to the normal distribution is also investigated. The class of statistics is assumed to be generated by a score function having two bounded derivatives and the rates are of the type $n^{-\frac{1}{2}}(\log n)$.
- 6. A robust spherical correlation coefficient against scale. There are several bi-directional situations where it is required to obtain a measure of correlation. The question has been raised often that the bi-directional correlation coefficient so far known are not scale invariant even asymptotically in the sense that their asymptotic distributions under the hypothesis of independence with von Mises marginals depend upon the concentration parameters. Following a conditional approach of Cox (J. Roy Statis. Soc. (1975) 380-381), a new correlation coefficient is introduced which is asymptotically robust in this sense. The asymptotic distribution under the hypothesis of independence is obtained and its properties are examined. An example is given for illustration.
- 7. Realization of &p by spaces of random variables. Let $L^0(\Omega, P)$ be the space of equivalent classes of random variables defined on the probability space (Ω, A, P) . Let # be the closed subspace of $L^0(\Omega, A, P)$ spanned by a sequence of i.i.d. (independent and identically distributed) random variables having the symmetric nondegenerate law F. It is proved that # is linearly homeomorphic to ℓ_p for 0 if F belongs to the domain of normal attraction of symmetric stable law with exponent p.

ALE FORTE OFFICE OF SCIENTIFIC RESEARCH (AFSC) NOTICE OF PRANCHITAT TO DDG This tachnocal reserve mas been reviewed and is approved for public relies a LAW AFR 190-12 7b). Distribution is unitable.

A. D. BLOSE Technical Information Officer

- Analysis of Central Place Theory. Central Place Theory predicts a regular spatial pattern in the plane, and we observe that the Delaunay triangles will be equilateral under the theory. However, when the pattern is 'random', Miles (Mathematical Biosciences (1970), 85-127), has given the asymptotic p.d.f. of the interior angles of a random Delaunay triangle. We propose a von Mises-type model with a concentration parameter κ ; the larger the value of K , the closer we are to the Miles' density for some value of K . We provide the moment and maximum likelihood estimators of κ , and it is recognized that the areas of the Delaunay triangles play an important role. We construct a test of departure from the random pattern with the alternative of Central Place Theory. As a numerical example, we analyze 44 Central Places in Iowa where we find some evidence for the validity of Central Place Theory in that particular region.
- 9. Sample size, parameter rates and contiguity -- the i.i.d. case. Let $\{\alpha_n\}$, $n \ge 1$, be a nondecreasing sequence of positive integers tending to $+\infty$ as $n \to +\infty$, and let $X_1, \ldots, X_{\alpha_n}$ be independent and identically distributed random variables defined on the probability space (X,A,P_A) , $\theta \in \Theta$, where Θ is an open subset of R^k , $k \geq 1$. Also, consider a sequence $\{\tau_n^{}\}$, $n \geq 1$, of positive real numbers tending to $+\infty$ as $n \to +\infty$ and set $\theta_{\tau_n} = \theta_n^{\tau} = \theta + h\tau_n^{-\frac{1}{2}}$, $h \in R^k$, $h \neq 0$. Let $A_{\alpha_n} = \sigma(x_1, \dots, x_{\alpha_n})$ be the σ -field generated by the random variables X_1, \dots, X_{α_n} . Denote the restrictions of the probability measures $~\textbf{P}_{\theta}~$ and $~\textbf{P}_{\theta}\,\tau~$ to the $\sigma\text{-field}$ A_{α_n} by $P_{n,\,\theta}^{\alpha}$ and $P_{n,\,\theta}^{\alpha,\,\tau}$, respectively. It is well known that the choice $\theta_n = \theta + hn^{-\frac{1}{2}}$, $h \in \mathbb{R}^k$, for the "moving" parameter θ_n , is sufficient to yield contiguous sequences of probability measures under mild conditions [see Roussas, <u>Contiguity of probability measures</u>: <u>some applications in statistics</u>, <u>Chapter 2</u>, especially p. 65, <u>Cambridge Univ. Press</u>, <u>London</u>, 1972]. In the present paper, the following question is investigated: Under what conditions on the sample size α_n and the parameter rate τ_n are the sequences of probability measures $\{p_{n,\,\theta}^{\alpha}\}$ and $\{p_{n,\,\theta}^{\alpha}\}$ contiguous? After establishing some auxiliary results, we prove that under suitable conditions, for each $\theta \in \Theta$, the sequences of probability measures $\{P_{n,\,\theta}^{\alpha}\}$ and $\{P_{n,\,\theta}^{\alpha,\,\tau}\}$ are contiguous if and only if $\alpha_n = O(\tau_n)$.

- Shorted operators and generalized inverses of matrices. Shorted operators were introduced by W.N. Anderson, Jr. [SIAM J. Appl. Math [20] 1971), 520-525] in connection with problems in electrical network analysis. The shorted operator of the positive definite operator A is the supremum over all positive operators of those which are smaller than A and which have a fixed subspace for their range. In the present paper, we give two explicit representations for the shorted operator, one in terms of the g-inverse, the other in terms of the least squares inverse of a complex matrix. In addition, some new theorems and characterizations about shorted operators are given. These results create a new role for the shorted operator which has found applications in several areas.
- A simple test for goodness-of-fit based on spacing with some efficiency comparisons. Let x_1, \dots, x_{n-1} be inde-11. pendent and identically distributed random variables with a common distribution function (df)G(x). Then, for the goodness-of-fit problem, i.e. the problem of testing that G(x) is equal to a specified continuous $df G_{0}(x)$, a class R_n based on spacings is proposed. of tests The exact as well as the asymptotic distribution of the proposed test statistic is derived. The Pitman asymptotic relative efficiencies and the Bahadur efficiencies of the proposed tests relative to some of the existing tests in literature are studied. Finally, the statistic R_n^* which has the maximum efficacy in the class of tests R_n , is investigated, and a table, to obtain the critical values of R_n^* is provided.
- 12. Local maxima of the sample functions of the N-parameter

 Bessel process. It is proved that almost every sample function of the N-parameter Bessel process has a local maxima.
 In addition, some properties related to the local maxima are
 investigated.
- 13. Matrices G satisfying simultaneous equations $A^*MAG = A^*M$ and $G^*NGA = G^*N$. For a matrix A, let M(A) denote its column span. Let $A_{(MN)}$ represent a matrix G such that $AG \in \{P_A\}$ and $GA \in \{P_G\}$ where P_A and P_G are projections into M(A) and M(G) under seminorms induced by M and N respectively. It is well known that when M and N are positive definite, $A_{(MN)}$ is the unique minimum N norm M least squares inverse (Moore Penrose inverse) of A. This however is no longer true when M and N are only positive semidefinite. What is true is that

$$\{A_{MN}\} \subset \{A_{(MN)}\} \tag{1}$$

(Theorem 3.2. in S.K. Mitra and C.R. Rao (1974). Linear Algebra and its Applications, 9, 155-167). The object of the present paper is to determine just how large is $\{A_{(MN)}\}$ compared to $\{A_{MN}\}$. A general solution to $A_{(MN)}$ is obtained along with the necessary and sufficient conditions for equality to hold in (i).

- Empirical Distribution Functions and Functions of Order Statistics Under Mixing. In this paper some fundamental properties of empirical distribution functions are derived in the case of mixing random variables. These results are then utilized to study asymptotic normality and strong law of large numbers for functions of order statistics. It is also shown that both the "linear in probability" bounds and the "almost sure nearly linear" bounds of Shorack (1972, Ann. Math. Statist., 412-427) and Wellner (1977, Ann. Statist., 473-480) respectively continue to be valid if the stationary sequence of random variables is φ-mixing. Some other bounds in probability are obtained which account for the strength of mixing when the Shorack-Wellner bounding curves do not hold.
- 15. Rank order estimates in the case of grouped data (linear regression). Let $\hat{\alpha}'$ and $\hat{\beta}'$ be the rank order estimators of the location and scale parameters respectively in the simple linear regression model. Let $\hat{\alpha}$ and $\hat{\beta}$ be their counterparts when measurements are rounded off, to say, the nearest integers with ties handled by the average scores method. Approximate large sample distributions of $\hat{\alpha}$ and $\hat{\beta}$ are obtained. Simple counter examples are given to show $\hat{\alpha}$ and $\hat{\beta}$ may fail to be approximately normal. The difference $|\hat{\beta}' \hat{\beta}|$ can be made negligible by the proper choice of regression constants. A procedure for computing $\hat{\alpha}$ and $\hat{\beta}$, as well as applications to Chemical Engineering and Education (wherein ties abound) are given.
- 16. Asymptotic normality and convergence rates of linear rank statistics under alternatives. Asymptotic normality of linear rank statistics under alternatives is proved employing techniques of Chernoff and Savage ((1958), Ann. Math. Statist., 29, 972-996). Under suitable assumptions the rate of convergence to normality under alternatives is also obtained for such statistics. The results of the paper are related to those of Hájek ((1968), Ann. Math. Statist. 39, 325-346) and Hoéffding ((1973), Ann. Statist., 1, 56-66). Results on the rate of convergence extend those of Jurecková and Puri ((1975), Ann. Prob. 3, 526-573) and Bergström and Puri ((1976), Ann. Statist. 671-680).

- 17. Invariance principles for rank statistics for testing independence. In this paper invariance principles are developed for rank statistics for testing independence, and suitable order (a.s.) of remainder term is found. The results obtained have useful applications to sequential tests for independence and tests of power one.
- 18. Asymptotic multinormality and remainder terms of linear rank vectors under alternatives. Asymptotic multinormality of linear rank statistics based on independent vector valued random variables is obtained. Under suitable assumptions, weak estimates for the remainder terms for convergence to normality are also obtained.
- Asymptotic normality of linear rank statistics with discontinuous score generating functions. Let X_{N1},...,X_{NN}, $N \ge 1$, be independent random variables with continuous distribution functions $F_{N1},\dots,F_{NN};$ and let R_{N1}^+ be the rank of $|X_{N1}|$ among $|X_{N1}|, \dots, |X_{NN}|$. Under different sets of conditions, we derive the asymptotic normality of simple linear signed rank statistics $S_{N}^{+} = \sum_{i=1}^{N} c_{Ni} a_{N} (R_{Ni}^{+}) \operatorname{sgn} X_{Ni}$ where c_{N1}, \ldots, c_{NN} are known regression constants, sgn x = 1 if $x \ge 0$, sgn x = -1 if x < 0, and are scores generated by a function $\psi(t)$, $a_{N(1)}, \dots, a_{N(N)}$ 0 < t < 1 which in contradistinction to the earlier literature is no longer assumed to be continuous. The results obtained are generalizations of the earlier results on limit theorems due to Hajek (1968, Ann. Math. Statist. 325-346) and Hušková (1970, Z. Wahrscheinlichkeitstheorie. Verw. Geb., 308-322), among others.
- 20. A nonparametric test for equality against ordered alternatives in the case of skewed data with a biomedical application. For testing the equality of k-samples against ordered alternatives, a new nonparametric test is proposed based on the ideas of Gastwirth (1965, JASA, 1127-1141), Hogg et al. (1975, JASA, 656-661) and Puri (1965, Communications in Pure and Applied Math. 51-63). The test is shown to be superior to the Jonckheere (1945, Biometrika, 135-145) test in the case of skewed distributions. An application to lung cancer data illustrates the theory.

- 21. Limit theorems related to distributions of certain linear rank statistics with stochastic predictors. Let $\{x_{ni}, 1 \le i \le n \ , \ n \ge 1\}$ be independent random variables with continuous cumulative distribution functions $\{F_{ni}, 1 \le i \le n \ , \ n \ge 1\}$ and let $\{R_{ni}, 1 \le i \le n \ , \ n \ge 1\}$ be their corresponding ranks. Consider a stochastic linear rank statistics $T_n = n^{-\frac{1}{2}} \sum_{i=1}^{n} c_{ni} \phi_{ni} \frac{R_{ni}}{n+1}$ where (c_{n1}, \ldots, c_{nn}) are random variables, not necessarily independent, and ϕ a finite function on (0,1). Under suitable assumptions, it is shown that the limit distribution of T_n is weighted normal which degenerates to the normal in extreme cases. These results are of importance in several problems of the analysis of variance and provide new insight when the standard normality assumptions do not hold.
- 22. On the degeneration of the variance in the asymptotic normality of signed rank statistics. For each $N \ge 1$, let X_{N1} , $1 \le i \le N$, be a sequence of independent random variables with continuous distribution functions F_{Ni} , $1 \le i \le N$, respectively; and let R_{Ni}^+ be the rank of $|X_{Ni}|$ among $|X_{N1}|, \ldots, |X_{NN}|$. Consider the statistic $S_N^+ = \int\limits_{i=1}^N c_{Ni} a_N(R_{Ni}^+) \operatorname{sgn} X_i$, where c_{Ni} , $1 \le i \le N$ are known regression constants, $\operatorname{sgn} x = 1$ or -1 according as $x \ge 0$ or <0, and $a_N(i)$, $1 \le i \le N$ are scores generated by some function $\psi(t)$, 0 < t < 1, either by $a_N(i) = \psi(i/(N+1))$, $1 \le i \le N$, or by a more general procedure. In this paper we investigate the asymptotic behavior of S_N^+ when $\operatorname{Var} S_N^+$ compared with the variance of S_N^+ under the hypothesis of symmetry is allowed some degree of degeneracy. For discrete score functions, under suitable assumptions on the c_{Ni} and the distribution functions, it is shown that if

the ratio $\text{Var }S_N^+/\sum\limits_{i=1}^N c_{Ni}^2$ goes to zero at most at the rate $N^{-\alpha}$, for some $0 < \alpha < \frac{1}{2}$, then S_N^+ is asymptotic normal with natural parameters $(\text{ES}_N^+, \text{Var }S_N^+)$ as well as with simpler parameters $(\mu_N^+, \frac{2}{3})$. An application to the study of divergent shift alternatives is given.

- 23. Stochastic integrals and rank statistics. In this paper, convergence of a class of stochastic integrals of weighted empirical processes is obtained. This result is in turn used to establish the result of Hájek (1968, Ann. Math. Statist. 39, 325-346), concerning the asymptotic normality of a simple linear rank statistic $S_{\rm N}$ under alternatives, but with relaxed regularity conditions on the regression constants. A centering constant $\mu_{\rm n}$ is obtained and the asymptotic degeneracy of $S_{\rm N}$ is discussed.
- 24. A note on predicting the results of chess championship.

 In this paper, the method of prediction of results of chess championship matches suggested by Saaty and Vargas (Behavioral Science, 1980, 25, 180-191) is examined. It is shown that the claimed accuracy of this method results partly from arithmetic errors, and, in general, it is impossible to achieve this accuracy with any reasonably high confidence
- Rank procedures for testing sub-hypotheses in linear regression. In the linear regression model $X_i = \alpha + \beta c_i + Z_i$ the problem of testing the subhypothesis that some (but not all) components of β are equal to 0 is considered in this paper. Two classes of asymptotically distribution-free tests for this problem are studied, one based on a quadratic form in aligned rank statistics and the other based on a quadratic form in rank-order estimates of regression parameters. We derive the asymptotic relative efficiencies of these rank procedures with respect to two classical parametric procedures: the general likelihood ratio test and the test based on least-squares estimates of regression parameters. Asymptotic optimality (à la Wald) is also discussed.
- 26. A class of nonlinear admissible estimators in the one parameter exponential family. The paper deals with the admissibility of nonlinear estimators of the form (aX+b)/(cX+d) in the one parameter exponential family, in estimating g(θ) with quadratic loss. The method employed is reminiscent of that of Karlin (1958, Ann. Math. Statist., 29, 406-436) who gave sufficient conditions for admissibility of linear estimators aX in estimating the mean in the one parameter family. Particular cases of estimators of the form c/X are studied and several examples are given. The case of truncated parameter space is also discussed.

- 27. Shorted matrices an extended concept and some applications in design and analysis of experiments. The concept of a shorted operator on the cone \mathcal{C}_{n} of nonnegative definite matrices of order n x n introduced by Krein and studied recently by Anderson and Trapp (SIAM Journ. Appl. Math. 1975) is extended to a wider class of matrices. For matrices in $\,\mathcal{C}_{_{_{\!\boldsymbol{n}}}}\,$, the shorting operation is permissible with reference to any subspace S of the n-dimensional Euclidean space E^n , provided restrictions are symmetrically placed on the row and column spans. For general matrices, the shorting operation is uniquely defined if and only if certain conditions are satisfied by the matrix itself and by subspaces providing restrictions on the row and column spans. The key point in this development is a theorem of Anderson and Trapp which exhibits the shorted n.n.d. matrix as the limit of a sequence of parallel sum matrices. It is also shown that the generalized shorted matrix is related to a g-inverse resembling the minimum semi-norm g-inverse in much the same way as the Krein-Anderson-Trapp shorted operator was shown to be related to the minimum semi-norm g-inverse itself in Mitra-Puri (Linear Algebra and its Applications, 1979, 25, 45-56). Some applications of the shorted operator in mathematical statistics with special reference to the design and analysis of experiments are provided.
- 28. The Fuzzy Integral. In this paper, the fuzzy integral of a positive, measurable function with respect to a fuzzy measure is defined. The monotone convergence theorem and Fatou's lemma are shown to be valid in the new setting. Some of the properties of the new integral are studied. A theorem stronger than the Lebesgue-dominated theorem is derived.
- 29. Rank order tests for the parallelism of several regression surfaces . For testing the hypothesis that several (s $_{2}$ 2) linear regression surfaces $x_{ki} = \alpha_{k} + \beta_{k} c_{ki} + z_{ki}$ (k = 1,...,s) are parallel to one another, i.e., $\beta_{1} = \ldots = \beta_{s}$, a class of rank-order tests are considered. The tests are shown to be asymptotically distribution-free, and their asymptotic efficiency relative to the general likelihood ratio test is derived. Asymptotic optimality in the sense of Wald is also discussed.
- 30. Testing for the parallelism of regression lines using rounded-off data, and a biomedical application. For testing for the parallelism of several regression lines, the existing non-parametric literature deals with the cases when the underlying observations originate from absolutely continuous distribution functions. The present paper drops this unrealistic

assumption and deals with a class of rank order procedures as applied to the case of rounded-off data, with ties handled by the averaged scores method. The results obtained are the strongest in this direction. An application related to Physical Fitness data is provided.

- In this paper we derive the log likelihood function for point processes in terms of their stochastic intensities, using the martingale approach. For practical purposes we work with an approximate log likelihood function which is shown to possess the usual asymptotic properties of a log likelihood function. The resulting estimates are strongly consistent and asymptotically normal (under some regularity conditions). As a by-product, a strong law of large numbers and a central limit theorem for continuous martingale are derived.
- 32. The Hausdorff α -dimensional measures of the level sets and the graph of the N-parameter Wiener process. Let $w^{(N,d)}$ be the N-parameter Wiener process with values in R^d . Then we prove that the dimension of the level sets is N-d/2 with positive probability if d < 2N. The dimension of the graph is a.s. $\min\{2N, N+d/2\}$. The level sets have zero (N-d/2)-measures a.s. and the graph has zero $\min\{2N, N+d/2\}$ -measure a.s.
- 33. Theory of nonparametric statistics for rounded-off data with applications. This 85-page article gives a unified account of nonparametric statistics, covering testing, estimation, multiple comparisons, analysis of variance and regression, for rounded-off data, with ties handled by the average scores method. Numerous applications to the Biomedical, Physical, Engineering and Behavioral Sciences are illustrated.
- 34. Optimal non-negative interpolating cubic splines. In this paper a new approach is provided to the construction of non-negative interpolating cubic splines. One considers first the optimal natural interpolating cubic spline which satisfies only the obvious constraint of passing through the data points. In each sub-interval between two successive knots in which this spline assumes negative values, this portion of the spline is replaced by a non-negative cubic spline. This is carried out by the addition of extra knots, the locations of which are variable depending upon the values of the original spline and its first two derivatives at the two end points of each sub-interval. In general five extra knots are required to determine a non-negative spline with the first two derivatives matching the original spline at the end

points and which is tangent to the abscissa at the middle extra knots. The spline constructed in this local manner is clearly a non-negative interpolating spline. Precisely the optimal non-negative interpolating spline is obtained in a local manner by the proper modification of the optimal unconstrained spline in those subintervals in which latter spline has negative excursions. (This is the first part of the paper. It is intended to derive an algorithm suitable for implementation on a digital computer).

- 35. Some remarks on strategy in playing tennis. The paper shows how the probability of winning a game, a set and a match, depend on the probability of winning a ball (from one's own and from the opponent's serve). The conditions are given under which the usual serve strategy (strong-weak) is optimal. It is also shown that a game is "strategy-less" in the sense that if the player has k (= 1,2 or 3) balls which he can play with increased probability of winning, then all strategies of placing these balls during the game are equivalent. Finally, it is shown that the last property does not carry over to the case of optimal strategy in a match.
- 36. On estimating intersubject variability of choice probabilities under observability constraints. This paper provides methods for estimating inter-subject variability of the probability of a given event defined in terms of subject's behavior (e.g. probability of a given choice in discrimination experiment). The constraints consist of using no more than two independent observations for each subject. Estimators are provided for assessing the inter-subject "variance" of the analyzed probabilities; also, a test variable is given for testing the hypothesis that the average probability is the same for two groups of subjects.

B. Ph.D. Dissertations under the direction of Professor M.L. Puri

1. Richard Michael Rubison: (i) Robustness of optimal polynomial regression lines. (ii) A class of two-sample tests for location and scale.

Problem 1. Robustness of optimal regression designs. The classical theory of D-, G-, a- and A- optimality studied by Kiefer, Wolfowitz, Elfving, Studden and Federov among others is discussed with respect to robustness. It is noticed that the optimal designs corresponding to these optimality criteria are somewhat "fragile" against misspecifications of the model, e.g., for polynomial regression of degree k these optimal designs are concentrated on (k+1) points and thus are inadequate for detecting that the true regression polynomial is of degree > k.

In contrast to the above optimal designs it is noted that the C-restricted D-optimal designs introduced by Stigler (1971, <u>JASA</u>, 66, 311-318) are robust against model misspecificiations.

Our approach to the robustness of these optimal polynomial designs is through the theory of moments and orthogonal polynomials. In this setting the various optimality criteria are examined.

<u>Problem 2. A class of two-sample tests for location and scale.</u> We consider a class of two-sample rank order tests for location and scale. This class contains as special cases the Lepage (1971, <u>Biometrika</u>, 60, 113-116) test which is a combination of the Wilcoxon and Ansari-Bradley statistics and several other tests including the normal scores type test obtained by combining the Van der Waerden and the Klotz statistics. The small as well as large sample properties of the proposed test statistics are studied and for the total sample size < 25, the critical values of the normal scores type test are tabled.

Navaratna S. Rajaram. Stochastic integrals of weighted empirical processes and an application to the limiting distribution of linear rank statistics. Let X_{N1},...,X_{NN} be independent random variables with continuous cumulative distribution functions F_{N1},...,F_{NN} respectively. Consider the rank order statistics

$$S_{N} = \sum_{i=1}^{N} c_{Ni} a_{N} (R_{Ni}) ,$$

where R_{N1}, \dots, R_{NN} are the respective ranks of X_{N1}, \dots, X_{NN} , C_{N1}, \dots, C_{NN} are known constants, and $a_{N(1)}, \dots, a_{N(N)}$ are

scores generated by a known function $\phi\left(t\right)$, 0 < t < 1 in a specified manner. Asymptotic normality of S_{N} is obtained by expressing it as the stochastic integral of suitably defined empirical and weighted empirical processes, and using techniques of convergence of stochastic processes and reproducing Kernel Hilbert spaces.

An alternative approach to the same problem is provided by using stability results from classical probability theory. A multivariate extension is also given.

The phenomenon of Gaussian noise is examined by viewing it is as the limiting action of a sequence of disturbances caused by finitely many random variables acting on a certain function space. A central limit theorem for white noise is proved. A connection between such noise phenomenon and the asymptotic behavior of certain rank statistics is exposed.

3. Ching-Yuan Chiang. Some rank order tests in linear regression. This thesis investigates the rank order tests of some hypotheses in multiple linear regression models. Consider the regression model $X_i = g + gc_i + Z_i$ (i = 1,...,n), where $_{\alpha}$ (the intercept) and $_{\xi}$ = $(\beta_{1},\ldots,\beta_{q})$ (the regression parameters) are unknown, c_1, \ldots, c_n are q-dimensional (column) vectors of known regression constants, and $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ are independent random variables with a common unknown continuous distribution function F(x). Let $\beta = (\beta_1, \beta_2)$, where $\beta_1 = (\beta_1, \dots, \beta_r)$, $\beta_2 = (\beta_{r+1}, \dots, \beta_q)$, $1 \le r < q$. We consider the following null hypotheses: (1) $H_1 : \beta_2 = 0$; (2) $H_2 : \alpha = 0$; (3) $H_3 : (\alpha, \beta_1) = 0$. For testing the null hypotheses H_1 , H_2 , and H_3 , we first estimate the nuisance parameters β_1 , β and β_2 , respectively, based on the ranks of the X_i 's, and then consider aligned rank order tests. For each problem, under suitable assumptions on the c_i 's and F , we derive the asymptotic distribution of the rank test statistics, which is central chi-square under the null hypothesis and non-central chisquare under a sequence of local alternatives, with the appropriate degrees of freedom. In each case we also derive the asymptotic relative efficiency of the rank tests with respect to a parametric competitor, the likelihood ratio test, which is asymptotically Wald-optimal. With a particular choice of scores, the proposed rank tests are asymptotically power-equivalent to the likelihood ratio tests. We also consider s (≥ 2) regression models

we also consider s (22) regression models $X_{i}^{(k)} = \alpha^{(k)} + \beta^{(k)} c_{i}^{(k)} + Z_{i}^{(k)} \quad (i = 1, ..., n_{k}, k = 1, ..., s) ,$

where the $\alpha^{(k)}$'s and the $\beta^{(k)}$'s are unknown parameters, the $c_i^{(k)}$'s are q-dimensional vectors of known regression

constants, and the $Z_1^{(k)}$'s are independent random variables with a common unknown continuous distribution. Then for the problem of testing the null hypothesis that the s regression surfaces are parallel to one another, we propose a class of aligned rank order tests. Under suitable assumptions, the proposed rank test statistics are shown to be asymptotically central chi-square under the null hypothesis and non-central chi-square under a sequence of local alternatives. The asymptotic optimality of the proposed rank tests is also studied.

4. Dan A. Ralescu. Admissibility of estimators in the one parameter exponential family and in Multivariate Location problems. Let X be a random variable with probability density function $f_{\theta}(\mathbf{x}) = f(\theta)e^{\theta \mathbf{x}}$ with respect to some ofinite measure \mathbf{u} ; $\theta \in \mathbf{0} = \{\theta : \beta^{-1}(\theta) = \int e^{\theta \mathbf{x}} d\mathbf{u}(\mathbf{x}) < \infty \}$. Sufficient conditions are obtained for the admissibility of nonlinear estimators of the form $\delta(\mathbf{x}) = (a\mathbf{x} + B)/(c\mathbf{x} + d)$ for the problem of estimating an arbitrary piecewise continuous, locally integrable function $g(\theta)$ with squared error loss. Several applications of the main result are given and some new nonlinear admissible estimators are discovered. The problem is also studied for the case when the parameter space is truncated, that is, when $\theta \in \theta_0 = \{\theta : \theta \leq \theta_0\} < \theta$, θ_0 being an interior point of θ . The minimaxity of linear estimators of the form $a\mathbf{x} + B$ in estimating an arbitrary differentiable function $g(\theta)$ is also investigated.

In the multivariate case the classes of estimators which improve upon the best invariant estimators are considered. Let X be a p-dimensional random vector with density $f(x-\theta)$ where $\theta \in \mathbb{R}^p$ is a location parameter. For $p \geq 3$, different classes of estimators $\delta(X)$ which are uniformly better than the best invariant estimator $\delta(X) = X$ are obtained when the loss function is of the type

$$L(\theta, \delta(\mathbf{x})) = \sum_{i=1}^{p} c_i(\theta_i - \delta_i(\mathbf{x}))^2$$

where c_1,\dots,c_p are given positive constants. It is shown that $\ell=\delta_1$ * ξ , where δ_1 is an estimator which improves upon δ_0 outside of a compact set, ξ is a suitable probability density in \mathbb{R}^p , and * denotes the convolution. Some examples of densities ξ (such as truncated densities) which generate estimators which improve upon δ_0 are given, and some problems of further research interest are also stated.

- 5. Stefan Ralescu. Asymptotic theory of rank order statistics and the problem of the rate of convergence in the Central Limit Theorem. For each $N \geq 1$, let X_{Ni} , $1 \leq i \leq N$ be a sequence of independent random variables with continuous distribution functions F_{Ni} , $1 \leq i \leq N$, and let R_{Ni}^+ be the rank of $|X_{Ni}|$ among $|X_{N1}|,\ldots,|X_{NN}|$. We are concerned with the asymptotic distribution of the statistic $S_N^+ = \sum\limits_{i=1}^N c_{Ni} a_N(R_{Ni}^+) sgn(X_{Ni}^-)$, where c_{Ni}^- , $1 \leq i \leq N$ are known (regression) constants, sgn = 1 or -1 according to $x \geq 0$ or < 0 , and $a_N(i)$, $1 \leq i \leq N$ are scores generated by a function $\psi(t)$, 0 < t < 1 , either by $a_N(i) = \psi(i/(N+1))$, $1 \leq i \leq N$, or by a more general procedure.
 - (a) For discontinuous score function ψ , under different sets of conditions on c_{Ni} 's , F_{Ni} 's and the nondegeneration of Var S_N^+ , the asymptotic normality of S_N^+ is derived with parameters (μ_N^+, σ_N^2) which are easier to handle in practical situations than $(ES_N^+, Var S_N^+)$. The results are obtained using the projection method together with a separate study of the score function which has just one jump and is constant otherwise.
 - (b) The asymptotic behavior of S_N^+ when $\text{Var } S_N^+$ under the hypothesis of symmetry is allowed some degree of degeneracy studied. For discrete score functions, under suitable conditions on the distribution functions and regression constants,

it is shown that if the ratio $\text{Var}(s_N^+)/\sum\limits_{i=1}^N c_{Ni}^2$ goes to zero at most at the rate $N^{-\alpha}$, for some $0<\alpha<\frac{1}{2}$, then s_N^+ is asymptotically normal with parameters $(\text{Es}_N^+,\text{Var}\ s_N^+)$ as well as with some other parameters (μ_N^+,σ_N^2) . An application to the study of divergent shift alternatives is given.

(c) Under suitable assumptions concerning the smoothness and the order of magnitude of the (unbounded) score function ψ near 0 and 1, the problem of the rate of convergence in the C.L.T. for S_N^+ (with $c_{N1}^-=1,\ 1\le i\le N)$ is investigated. If ψ increases to infinity more rapidly than $-\log(1-t)$ as $t\ne 1$, the order of normal approximation is found to be $O(N^{-\alpha})$ for some $0<\alpha<\frac{1}{2}$. The proof relies on certain properties of empirical distribution functions in the non-i.i.d. case which are of independent interest.

- C. Publications (including the papers submitted) under the present grant.
 - 1. Puri, M.L. "Augmenting Shapiro-Wilk test for normality." Contributions to Applied Statistics, Birkhauser Verlag (1976), 129-139. Co-author: C.R. Rao.
 - Puri, M.L. "On linear combination of order statistics." Essays in Probability and Statistics (1976), 433-449. Co-author: F. Eicker.
 - 3. Puri, M.L. "Problems of association for bivariate circular data and new test of independence." <u>Proceedings of the Fourth International Symposium on Multivariate Analysis</u>, North Holland Publishing Company (1977), 513-522. Co-author: J.S. Rao.
 - 4. Puri, M.L. "Asymptotic distribution-free aligned rank order tests for composite hypotheses for general linear models." Zeitschrift für Wahrscheinlichkeitstheorie und Verw. Geb., (1977), 39, 175-186. Co-author: P.K. Sen.
 - 5. Puri, M.L. "Convergence and remainder terms in linear rank statistics." Annals of Statistics, (1977) 5, 671-680. Co-author: H. Bergstrom.
 - 6. Puri, M.L. "A robust spherical correlation coefficient against scale." <u>Biometrika</u> (1978) 65, 391-395. Co-author: K.V. Mardia.
 - 7. Puri, M.L. "Asymptotic behavior of stochastic linear rank statistics." <u>Proceedings of the International Statistical Institute</u> (1979), 392-397. Co-author: N.S. Rajaram.
 - 8. Puri, M.L. "Realization of lp by spaces of random variables." <u>Applicable Analysis</u> (1979) <u>8</u>, 337-347. Co-author: S.S. Sheu.
 - 9. Puri, M.L. "Analysis of central place theory." <u>Proceedings</u> of the International Statistical Institute (1979), 93-110. Co-authors: R. Edwards and K.V. Mardia.
 - 10. Puri, M.L. "Sample size, parameter rates and contiguity -the i.i.d. case." <u>Communication in Statistics</u> -- <u>Theory</u>
 and <u>Methods</u>, (1979) A8(1), 71-83. Co-authors: M.G. Akritas
 and G.G. Roussas.
 - 11. Puri, M.L. "Shorted operators and generalized inverse of matrices." Journal of Linear Algebra and its Applications (1979) 25, 45-56. Co-author: S.K. Mitra.

- 12. Puri, M.L. "A simple test for goodness-of-fit based on spacing with some efficiency comparisons." Contributions to Statistics -- Jaroslav Hájek Memorial Volume, Academia Prague, Czechoslovak Academy of Sciences (1979), 197-209. Co-authors: J.S. Rao and Y. Yoon.
- 13. Puri M.L. "Local maxima of the sample functions of the N-parameter Bessel process." Stochastic Processes and Their Applications, 9 (1979), 137-145. Co-author: L.T. Tran.
- 14. Puri, M.L. "Matrices G satisfying simultaneous equations A*MAG=A*M and G*NGA=G*N." <u>Journal of the Indian Statistical Association</u> (1980), 103-108. Co-authors: C.G. Khatri and S.K. Mitra.
- 15. Puri, M.L. "Empirical distribution functions and functions of order statistics for mixing random variables." <u>Journal of Multivariate Analysis</u> (1980) <u>10</u>, 405-425. Co-author: L.T. Tran.
- 16. Puri, M.L. "Rank order estimates in the case of grouped data." Sankhyā, Series B (1980) 41, 239-259. Co-author: A.R. Padmanabhan.
- 17. Puri, M.L. "Asymptotic normality and convergence rates of linear rank statistics under alternatives." <u>Banach Center Publications</u> (1980) <u>6</u>, 267-277. Co-author: N.S. Rajaram.
- 18. Puri, M.L. "Invariance principles for rank statistics for testing independence." <u>Contributions to Probability Theory</u>. Academic Press, New York (1981), 267-282. Co-author: L.T. Tran.
- 19. Puri, M.L. "Asymptotic multinormality and remainder terms of linear rank vectors." (Submitted).
- 20. Puri, M.L. "A note on predicting the results of chess championship." Behavioral Science (1981), 26, 85-87. Co-author: Robert Bartoszyński.
- 21. Puri, M.L. "On the degeneration of variance in the asymptotic normality of signed rank statistics." <u>Statistics and Probability</u>: <u>Essays in Honor of C.R. Rao</u>. North Holland Publishing Company, (To appear in 1981). Co-author: Stefan Ralescu.
- 22. Puri, M.L. "A nonparametric test for equality against ordered alternative in the case of skewed data with a biomedical application." (To appear in 1981).
- 23. Puri, M.L. "Stochastic Integrals and Rank Statistics."

 <u>Proceedings of the International Colloquium on Nonparametric Statistical Inference</u>, North Holland Publishing Company.

 (To appear in 1981). Co-author: N.S. Rajaram.

- 24. Ralescu, Dan A. and Ralescu, Stefan: A class of nonlinear admissible estimators in the one parameter exponential family. Ann. Statist. 6 (1981), 177-183.
- 25. Ralescu, Dan A. The fuzzy integral. <u>Journal of Mathematical Analysis and Applications</u> 75, (1980), 562-570.
- 26. Puri, M.L. "Shorted matrices an extended concept and some applications." <u>Linear Algebra and its Applications</u> (accepted). Co-author: S.K. Mitra.
- 27. Puri, M.L. "A class of rank order tests for the parallelism of several regression surfaces." Annals of Statistics (Submitted).
- 28. Puri, M.L. Rank procedures for testing subhypotheses in linear regression." <u>Journal of Multivariate Analysis</u> (submitted). Co-author: Ching-Yuan Chiang.
- 29. Puri, M.L. "Maximum likelihood estimation for stationary point processes." <u>Annals of Probability</u> (Submitted).
- 30. Puri, M.L. "The Hausdorff α -dimensional measures of the level sets and the graph of the N-parameter Wiener process." (Submitted). Co-author: L.T. Tran.
- 31. Puri, M.L. "Theory of nonparametric statistics for rounded-off data with applications." (Submitted).
- 32. Puri, M.L. "Optimal non-negative interpolating cubic splines." (Submitted). Co-author: R.C. Davis.
- 33. Puri, M.L. "Some remarks on strategy in playing tennis." (Submitted). Co-author: Robert Bartoszyński.
- 34. Puri, M.L. "Linear models with stochastic predictors." (Submitted).
- 35. Puri, M.L. "On estimating intersubject variability of choice probabilities under observability constraints." (Submitted). Co-author: Robert Bartoszyński.
- 36. Puri, M.L. and Ralescu, Dan. "Improving upon the best invariant estimator in multivariate location problems." (Submitted).
- 37. Rubison, Michael Richard: (i) Robustness of optimal polynomial regression lines. (ii) A class of two sample tests for location and scale. Ph.D. Thesis (1976).
- 38. Rajaram, Navaratna S. Stochastic integrals of weighted empirical processes and an application to the limiting distribution of linear rank statistics. Ph.D. Thesis (1976).

- 39. Chiang, Ching-Yuan. Some rank order tests in linear regression. Ph.D. Thesis (1979).
- 40. Ralescu, Dan A. Admissibility of estimators in the one parameter exponential family and in multivariate location problems. Ph.D. Thesis (1980).
- 41. Ralescu, Stefan. Asymptotic theory of signed rank statistics and the problem of rate of convergence in the Central Limit Theorem. Ph.D. Thesis (1981).

D. Honors (from 1976 onward)

- Invited by the Stefan Banach International Mathematical Center, Warsaw, Poland, to give some lectures to the graduate and post-doctoral fellows delegated by the academies of Bulgaria, Czechoslovakia, German Democratic Republic, Hungary, Romania, and U.S.S.R.
- 2. Invited by the Commonwealth Scientific and Industrial Research Organization to visit some of its divisions of Mathematics and Statistics in Australia.
- 3. Invited by the University of California, Irvine, to give a a series of lectures during 1977-78 on a suitable topic in statistics under its "Distinguished Visitors Program".
- 4. Invited by the Executive Office of the President, Office of Management and Budget, Washington, D.C. to participate in the project, "Framework for Planning U.S. Federal Statistics 1978-89", by reviewing and commenting on materials of particular interest.
- 5. Invited by the Bernoulli Society for Mathematical Statistics and Probability to be one of the three co-editors of the two volume Proceedings of the International Colloquium on Non-parametric Statistical Inference. The other co-editors are Professors B.V. Gnedenko (Moscow State University) and Istvan Vincze (Hungarian Academy of Sciences). The Proceedings are being published by the North Holland Publishing Company, Amsterdam.

E. Professional Service (from 1976 onward)

- Member, Institute of Mathematical Statistics Committee on Summer Research Institutes, 1976, 1977, 1978, 1979, 1980, 1981.
- Member, Editorial Board of the International Journal, "Journal of Statistical Planning and Inference".
- 3. Member, Editorial Board of the International Journal, "Zeitschrift Mathematische Operationsforschung und Statistik".
- 4. Member, Nominating Committee of the Institute of Mathematical Statistics, 1977.
- 5. Participant, Visiting Lecturer Program in Statistics, sponsored by the Institute of Mathematical Statistics, the American Statistical Association, and the Biometric Society, 1977, 1978, 1979, 1980, 1981.

- 6. Reviewer, Research Proposals for the National Science Foundation and the National Research Council of Canada. Also, Reviewer of Proposals for Regional Conferences for the Conference Board of Mathematical Sciences.
- 7. Reviewer, Mathematical Reviews and Referee for several National and International Journals.
- 8. Member, International Organizing Committee, "Colloquium on Non-parametric Inference", sponsored by the Bernoulli Society for Mathematical Statistics and Probability, June, 1980, held in Hungary.
- F. Invited Addresses (from 1976 onward)

Professor M.L. Puri gave invited talks at the following places.

- I. International (outside the United States and Canada).
- 63rd Session of the Indian Science Congress, Waltair, India, "Linear models revisited", January 3-7, 1976.
- 2. Stephan Banach International Mathematical Center, Warsaw, Poland, three lectures on "Asymptotic methods in statistics", December, 1976.
- Freie Universitat Berlin, West Germany, two lectures:
 (a) "A unified theory of rank order estimation", (b) "Rank order test for composite hypotheses", December, 1976.
- 4. Universitat Koln, Koln, West Germany, "Asymptotic methods in general linear models", December, 1976.
- 5. Universite de Liege, Liege, Belgium, three lectures:
 (a) "A unified theory of rank order estimation I", (b) "A
 unified theory of rank order estimation II", (c) "Rank order
 tests for composite hypotheses", December 1976.
- 6. Technische Hochschule Aachen, West Germany, "Asymptotic methods in general linear models", December, 1976.
- 7. Eidg Technische Hochschule, Lausanne, Switzerland, "A unified theory of rank order estimation", December, 1976.
- University of Leeds, Leeds, England, "Unified theory of estimation based on rank statistics with applications", December, 1976.
- 9. University of Auckland, New Zealand, Visiting Professor, eight lectures on topics in Probability and Statistics, May 29-June 28.
- 10. University of Waikato, Hamilton, New Zealand, "Distribution-free procedures forgeneral linear models", June, 1977.

- 11. Commonwealth Scientific and Industrial Research Organization, Sydney, Australia, "Rank order tests for composite hypotheses", June 28-July 1, 1977.
- 12. Commonwealth Scientific and Industrial Research Organization and Australian National University, Canberra, Australia, "Distribution-free procedures for general linear models", July 1-5, 1977.
- 13. Commonwealth Scientific and Industrial Research Organization, Melbourne, Australia, "Testing against trend", July 5-8, 1977.
- 14. University of Melbourne, Australia, "Nonparametric methods in regression analysis", July 6, 1977.
- 15. Monash University, Melbourne, Australia, "Some aspects of nonparametric inference", July 7, 1977.
- 16. Forty-first Session of the International Statistical Institute, New Delhi, India, "analysis of central place theory", December 5-15, 1977.
- 17. Institute of Mathematical Statistics, New Delhi, India, "Rank order methods in regression analysis", December 16-18, 1977.
- 18. The International Conference in Statistics, Tokyo, Japan, "Linear models with stochastic predictors", November 28-30, 1979.
- 19. University of Patras, Patras, Greece, three lectures, December 3-6, 1979.
- 20. Colloquium on Nonparametric Statistical Inference (sponsored by the Bernoulli Society for Mathematical Statistics and Probability), Budapest, Hungary, June 23-28, 1980.
- II. Invited Addresses (within the United States and Canada)
 - 1. 1976-77

State University of New York, Stony Brook, University of Kentucky, Lexington; Kent State University, Kent Ohio; Bowling Green State University, Bowling Green, Ohio; McMaster University, Hamilton, Canada; University of Western Ontario, London, Canada.

2. 1977-78

University of California, Santa Barbara; University of Oklahoma (two lectures); State University of New York, Albsny, New York; University of California, Irvine (a series of lectures on Asymptotic Methods in Nonparametric Statistical Inference); Portland State University, Portland, Oregon; University of California, Los Angeles; The Institute of Mathematical Statistics meeting in New Delhi, India; University of California, Los Angeles; Portland State University, Portland.

3. 1978-79

University of Florida, Gainesville; Clemson University, Clemson, South Carolina; University of Maryland, College Park, Maryland, University of Massachusetts, Amherst, Massachusetts; University of Delaware, Newark, Delaware; Universityof Pittsburgh, Pittsburgh, Pennsylvania; Johns Hopkins University, Baltimore, Maryland.

4. 1979-80

University of Windsor, Windsor, Canada; Rice University, Houston, Texas; The Institute of Mathematical Statistics (Central Region Meeting), Iowa City, Iowa; International Symposium in Statistics and Related Topics, Ottawa, Canada.

G. Personnel Supported

- (1) M.L. Puri (Principal Investigator)
- (2) Richard Michael Rubison (Student Research Assistant)
- (3) N.S. Rajaram (Student Research Assistant)
- (4) Chiang-Yuan Chiang (Student Research Assistant)
- (5) Dan A.Ralescu (Student Research Assistant)
- (6) Stefan Ralescu (Student Research Assistant)

Rubison, Rajaram, Chiang and Dan Ralescu finished their degrees during the period of this grant. All of them (except Rajaram) are holding academic positions in the universities. Rajaram is a senior statistician at Lockheed (engineering and Management Services Company, Inc.) at Houston, Texas. Stefan Ralescu's thesis is almost completed. His final (defense of thesis) examination is March 12, 1981.

Respectfully submitted,

Anadom heri

Madan L. Puri Professor of Mathematics

KEY WORDS (Continue on reverse elde if necessary and identify by block number)
Circular data, nonparametric procedures, order statistics, tests of normality, Limit Theorems, generalized inverse of matrices, shorted operators, decision theory, admissibility, Wiener processes, Fuzzy Integrals, splines, Bessel Process, strategy.

20 ABSTRACT (Continue on reverse side if necessary and identify by block number)

This is the final report of the work done under Grant Number AFOSR 70-2927. deals with the research accomplished in the following areas: Analysis of circular data, order statistics, nonparametric procedures in testing as well as estimation, tests of normality, Limit Theorems, polynomial splines, generalized inverse of matrices and their applications in statistics, Wiener processes, game theory, Fuzzy integrals, decision theory and the problems of the speed of convergence.

DD 1 JAN 73 1473

402522 SECURITY CLASSIFICATION OF THIS PAGE (When Date Entended)

